DOCUMENT RESUME

ED 333 032 IR 015 025

AUTHOR Lee, William W.; Roadman, Kenneth H.
TITLE Ensuring Instructional Soundness through Sound Evaluation.
PUB DATE 90
NOTE 15p.; Paper presented at the International Conference of the Association for the Development of Computer-based Instructional Systems (32nd, San Diego, CA, October 28-November 1, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Formative Evaluation; Industrial Training; *Instructional Development; *Instructional Effectiveness; *Interactive Video; Material Development; Models; Needs Assessment; Postsecondary Education; Staff Development; *Summative Evaluation

ABSTRACT

Interactive videodisc training will only come into its own when its effectiveness for improving performance in the workplace is documented through carefully planned and implemented evaluation activities. This paper suggests a model for thorough, fool-proof evaluation to enter each phase of training course production and the rationale for each activity. The evaluation process has been separated into five distinct phases. These phases may not be as clearly differentiated in the production process used by individual companies but the intent is that all activities should be considered in order to tailor the final training product to the needs of the client. These five phases are: (1) Pre-contract; (2) Assessment/Analysis; (3) Design; (4) Development; and (5) Implementation. During the first phase, the client's expectations for evaluation are established and finalized so that a cost base for including evaluation in the overall product development can be established. The second phase is where needs assessment and front-end analysis are conducted. The third phase involves the establishment of the formative evaluation plan, validating the content, and preliminary work for summative evaluation. Evaluation during the fourth phase involves completion of the formative evaluation components in courseware production. The final phase involves the completion of the formative evaluation plan and conducting a summative evaluation. This paper shows that evaluation is not something to be tacked on after a course is developed. It must be woven throughout the entire production process. (9 references) (Author/BBM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED333862

ENSURING INSTRUCTIONAL SOUNDNESS THROUGH SOUND EVALUATION

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

BY

William W. Lee, Ph.D.
and
Kenneth H. Roadman, D.Ed.

Introduction

If any instructional delivery system is to become widely accepted, it must prove to be more effective than its predecessor. So it will be with Interactive Videodisc Instruction (IVI). As competition for contracts becomes more intense, IVI developers are beginning to realize that they need to be accountable for the skills their products claim to teach. Accountability means that interactive video courses will have to be more intensely scrutinized and evaluated. As Basarab (1989) aptly puts it "the value added to the credibility of training organizations by a standard education evaluation process is extensive." (p. 54)

Business and Industry are the largest consumers of IVI but show little interest in ensuring the soundness of instruction that they provide. These consumers must be convinced that the money spent evaluating training products is worth the expense. The primary reason for establishing instructional soundness is to justify the expense of development. Training must transfer into the workplace. This does not happen without encouragement from management. One way is to incorporate the skills from the training into performance appraisals.

Those courseware development companies that have discovered and have attempted to implement evaluation as part of their development process realize that evaluation is a complicated process. But, as Reeves (in press) has stated, design without evaluation is no more than blind trial and error creation.

This paper focuses on evaluation of IVI, however sound evaluation procedures are the same for any type of training course. The thrust here is to demonstrate that IVI will only come into its own when its effectiveness is thoroughly documented.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William W. Lee

Kenneth H. Roadman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

1201S025

Criteria for Evaluation

Evaluation terminology is used rather loosely and evaluation itself receives a great deal of "lip service" but there is little hard evidence of commitment to thorough evaluation. A notable exception is National Cash Register (Basarab, 1989). This is because of the parameters evaluation sets when installed within a preexisting courseware production process. Though the process seems complicated, evaluation need not be a looming monster. Thorough program planning and needs assessment can ease the pain of implementing evaluation.

Miller and Irving (1988) outline five criteria for preparing valid evaluations: 1) budget for formative and summative evaluation, 2) design evaluation early in program development, 3) use valid and reliable measures, 4) conduct evaluation with integrity, and 5) maintain communication with the client.

A Complete Evaluation Process

A thorough evaluation process that is constructed around principles of research can be designed to fit into either highly or loosely-structured models of courseware development. For our purposes, we will examine how evaluation can be integrated into a highly structured model. The steps for integrating evaluation into a more loosely structured model must be determined by the user of that model. Ultimately, all steps needed to achieve a desired level of evaluation must be determined during initial planning and analysis.

Defining Terms

Before examining the complete evaluation process, it is necessary to clearly define certain terms so that everyone speaks the same language and approaches evaluation from the same perspective. The client's understanding of the following terms is paramount to convincing him or her that evaluation pays off:

Formative Evaluation -- All activities that occur from the time a client begins contract negotiation until the final product is delivered which ensure the instructional soundness and suitability for use of an education and training program. The primary purpose of formative evaluation is product improvement.

Summative Evaluation -- The long-term tracking of an operating program which tests the effectiveness of the education or training program. The primary purpose of summative evaluation is to evaluate product effectiveness.

Validation -- Procedures employed to ensure that a training product teaches or tests what it claims to teach or test.

Validity -- A quantifying value used to describe the degree to which a training product teaches or tests what it claims to teach or test.

Reliability -- A quantifying value used to describe the degree to which a training product produces consistent results in what it teaches or tests.

Validity/Reliability Interaction -- A course may produce consistent results (reliability) when used to teach or test various users. However, the course may not be valid because it does not teach what it intends to teach. Reliability follows validity in that, if course objectives, content, and test items are congruent, the consistency of results achieved will determine that a user has learned or failed to learn the material on the basis of having taken the course or passed the tests. Validity rules out chance or incidental factors that might cause a user's score to reflect anything other than what was actually learned from the content of the training program. In essence, reliable measures may not be valid measures, but valid measures will always be reliable.

Instrument Validity -- Developing and evaluating test instruments to ensure that the informational data collected is unbiased and able to be replicated during subsequent administrations of the instrument.

Test Item Validity -- Statistical analysis of test items to ensure that they measure the skills learned to a sufficiently high enough degree to discriminate between high-achieving students and low-achieving students.

A Five Phase Project Evaluation Model

Table 1 lists the evaluation activities that are built into the five phases of product development. Specific activities that take place, their purpose, and the outcome are described in detail.

Phase I -- Pre-contract establishes a client's intended use for a training program and expectations regarding formative and summative evaluation. The client should be made aware that specific uses for a training program will require various levels of formative evaluation. At this stage it should be determined if performance appraisal is tied to the client's reward system. If a client intends to incorporate a training course into a job performance analysis, content validity and test item validity should be established to appropriately reward employees for their efforts. If the training is to be used only for self-improvement and has no bearing on subsequent job performance or career path, lower level options are required for evaluation. Table 2 outlines a series of if/then statements and the corresponding levels of evaluation.

TABLE 1
Evaluation Process

Phase	Activity	Purpose	Outcome
PHASE I *Pre-Contract*	• Establish client expectations regarding Formative and Summative Evaluation	• Provide information on evaluation	• Client knows that evaluation is critical to the successful implementation of a training program
	• Present Formative/ Summative Evaluation Documentation	• Inform client of the importance of the evaluation process and its relation to the delivery of a high quality, effective program • Inform client of available evaluation options	• Client fully understands the evaluation process: - Its importance in the delivery of a high quality, effective program - The client's role in providing necessary personnel to fully implement the evaluation process
	• Finalize client's expectations regarding the extent of evaluation	• Finalize costing of evaluation	• Final cost of evaluation • Client approves the scope of the evaluation
	• Establish Evaluation Timelines	• Provide focus for analysis activities.	• Program evaluation objectives
PHASE II *Assessment/ Analysis*	• Conduct Needs Assessment - Selection of analysis instruments - Collect and interpret analysis data	• Establish clarity of client need • Ensure appropriate data is collected • Collect and ensure accuracy of data collected	• Prioritized list of client needs • List of analysis instruments • Accurate data; unbiased interpretation
	• Conduct Front-end Analysis - Selection of analysis instruments - Collect and interpret analysis data - Establish readability level	• Designate/design those aspects of hardware/ software to collect data on-line	List of target audience characteristics.
	• Specify technical requirements for Formative and Summative Evaluation	• Provide focus for Design activities.	• Prioritized list of needs

TABLE 1
Evaluation Process (Continued)

Phase	Activity	Purpose	Outcome
PHASE III *Design*	• Formulate specifics of Formative and/or Summative Evaluation	• Establish the Formative Evaluation Plan	• Operational Formative/ Summative Evaluation Plan
		• Establish Summative Evaluation Plan; front-end considerations (establishing control groups, data collecting/ analysis procedures, subcontracts)	• Objectives appropriate for client's needs
			• Client signs off on: - prototype - storyboards - sample lessons
		• Select/develop appropriate forms and questionnaires	
		• Establish validation procedures for forms and questionnaires	
	• Validatecontent	• Ensure soundness of testing instruments	• Instructional Soundness
	• Validate readability of course content	• Ensure adherence to identified level of readability	
		• Provide focus for Development Activities	• Evaluation section of High Level Design Report
	• Review specifics of Formative/Summative Evaluation	• Finalize all elements formative and summative evaluation.	• Detailed design evaluation activities report
PHASE IV *Development*	• In-house review - Audio - Video - Graphics - Course Programming	• Major debugging	• Functioning course
		• Assurance that program changes are carried out	
	• Validate readability of course documentation	• Ensure adherence to identified level of readability	• Detailed Production Evaluation Activities Report

TABLE 1
Evaluation Process (Continued)

Phase	Activity	Purpose	Outcome
PHASE V *Implementation*	• Conduct Formative Evaluation - Administer Test - Conduct pilot study with typical user group - Collect and analyze data	• Validate Test • Improve product quality - Identify minor bugs - Identify areas of inconsistency and/or confusion	• Validtesting instruments • A product that runs or operates according to design specifications and is ready for full-scale implementation • Formative Evaluation Report
	• Conduct long-term Summative Evaluation	• Test the effectiveness of the product with respect to: - Achievement - User Characteristics - Improved job performance - Motivation - Behavior Change	• Data and information for client on the effectiveness of the product • Summative Evaluation Report

© 1991 by William W. Lee, Kenneth H. Roadman, Robert A. Mamone. Used with permission of the authors.

TABLE 2 Dependent Variables Determining Extent of Evaluation	
Variable	Evaluation
1. If the purpose of the course is only self-improvement...	Then establish face validity.
2. If the course will be marketed and/or 3. If the course will become part of a performance appraisal and/or 4. If student achievement will be tracked...	Then establish content validity.
5. If tests will be validated during a pilot study...	Then content validity should be established.

©1990 William W. Lee
Used with permission of the author.

The client receives a formal document, a presentation, or a combination of both, containing information about the importance of the evaluation process and its relation to the delivery of a high-quality product that optimally meets their training needs. It establishes the parameters to implement the proposed evaluation activities. Solidifying the extent of evaluation helps establish both the cost of the evaluation and timelines for implementing the necessary activities that occur during Assessment/Analysis, Design, Development, and Implementation. The deliverable is a list of program evaluation objectives that provides the focus for Assessment/Analysis activities.

Phase II -- Assessment/Analysis. "If anything deserves to be buried today, it is the technology persepctive that declares the solution before evaluating the problem." (Miller, 1989, p. 2). This statement poignantly and succinctly summarizes the importance of the Assessment/ Analysis Phase in the life of a project.
Assessment/Analysis is comprised of two basic processes, Needs Assessment and Front-End Analysis. Either or both may be implemented depending on client needs:

1. **Needs Assessment** is used to identify the client's training needs. Instruments, in the form of questionnaires, and procedures, such as mailings and telephone or on-site interviews, are established. Data are collected and interpreted resulting in a prioritized list of those needs.

Needs Assessment is a systematic process for determining goals, identifying discrepancies between goals and strategies, and establishing priorities for action. Needs Assessment can be done at the state level, national level, and the application to the course level. It should be the first step in a systematic approach to the development of instructional materials for a course or training program.

Types of needs include (Briggs, 1977):

Normative -- compared to a standard
Felt Need -- what people think they need
Expressed/Demand -- supply and demand
Comparative -- some people have it and some don't
Anticipated/Future Needs -- Projected demands

There are four steps to Needs Assessment:

1. Identify the broad range of goals.
2. Rank the goals in order of importance.

3. Identify discrepancies between expected and actual performance required to meet a goal.
4. Set priorities for action.

Many clients may have already conducted this Needs Assessment. If this is the case, their results should be provided to the instructional designer. Additional data from the Needs Assessment may be needed for clarification. With the client's own assessment completed, issues involving the delivery system for the intended training should be discussed. Conducting a Needs Assessment or reviewing the client's results will allow the developer to recommend the best delivery system. The results of Needs Assessment, or the client preferred delivery system, will be the guideline for all subsequent activities. However, the strengths and weaknesses of that system should be explained to prevent disagreement or dissatisfaction with the final product.

2. **Front-end Analysis** is the collection of all relevant data regarding a project and employs the tools of:

Audience Analysis -- identifies the learning characteristics of the potential users.

Critical Incident Analysis -- identifies the strengths and weaknesses of an existing training program.

Extant Data Analysis -- identifies existing training materials, manuals, references, and syllabi.

Goal Analysis -- transforms statements into tangible behaviors that serve as proof that a goal has been achieved. It should also describe how information will be processed and demonstrated.

Instructional Analysis -- identifies the subordinate or prerequisite skills that must be mastered for a student to achieve the instructional outcome. The skills may be psychomotor, intellectual, verbal, or attitudinal.

Subject Matter Analysis -- identifies the specific content related to a program, which enables the designer to develop content appropriate for the audience level.

Task Analysis -- a detailed description of the job-related tasks that the potential user must perform after completing the training.

Instruments for the assessment are selected, developed, and used to collect necessary information. The technical requirements for formative and summative evaluation will also be established to provide the focus of the Design activities regarding test development, data capturing, and on-line data reporting.

Phase III -- Design formulates the specifics of the formative and/or summative evaluation plans. The selection and development of appropriate survey instruments to establish the extent of validity begins here. Experimental and Control groups, data collection and analysis procedures, and subcontracting requirements establish appropriate objectives for the training program. The result is an operational Evaluation Plan. Where useful, a prototype, storyboards, and/or sample lessons are developed for client approval.

Measuring the validity and reliability of an IVI program is a complicated process. Often there are no tests or questionnaires available for highly technical content. Establishing the validity and reliability of a newly designed evaluation instrument must begin during this phase. Where measurement instruments do exist, they must be identified and determined to have undergone extensive prior validation and reliability testing to be certain that they measure the criteria they claim to assess. *Burros Mental Measurement Yearbook* is an excellent source of this information.

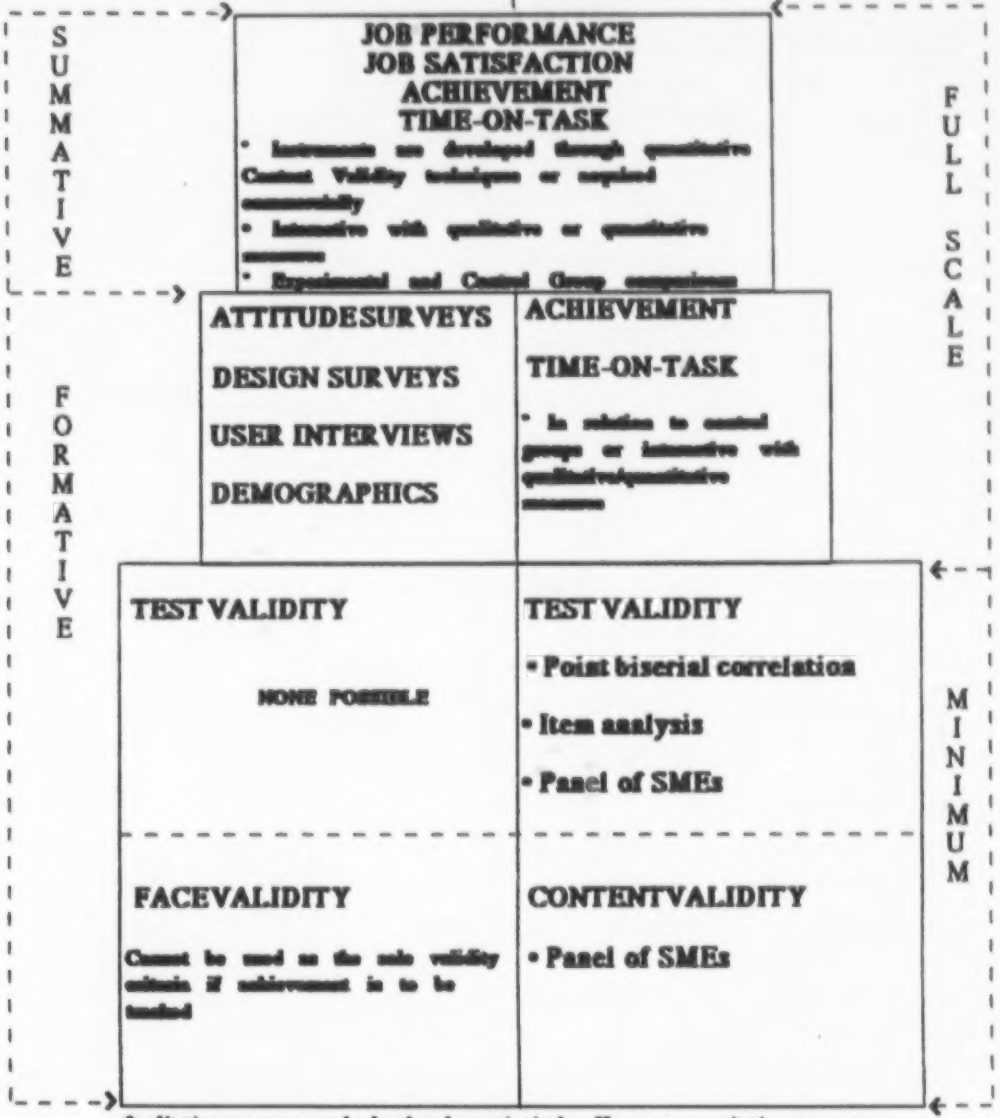
The level of validity agreed upon during the Pre-contract phase is finalized by the end of Design.

Table 3 represents a hierarchy of activities that lead to the highest levels of quality instruction and the relationship of quality to quantity elements of instructional soundness. Quantitative validity uses empirical data and results of statistical analyses to represent it while qualitative validity is assessed by analyzing more subjective characteristics. However, the methods for determining validity through qualitative means are no less effective than quantitative measures if instruments used to collect information have been tested through instrument validity.

The instructional soundness of a training program increases proportionally as each of four validity options are implemented. These four options are One-to-One, Small Group, Field Test, and Summative Evaluation. Each option is detailed below.

TABLE 3

Evaluation Activities Hierarchy



Qualitative measures can be developed quantitatively. However, quantitative measures are severely weakened if substituted by qualitative.

BEST COPY AVAILABLE

One-to-One -- An evaluation coordinator works with the Subject Matter Expert (SME) to establish the validity of the course or lesson content. This is done in conjunction with the final sign-off on storyboards and establishes that: 1) the objectives for the course match the content, and 2) test items match the objectives and are unambiguous. Two levels of validating the content are necessary depending on the client's purpose and intended use of the training, previously established in the Analysis Phase. These two levels are Face Validity and Content Validity.

Face Validity is the certification by a SME that the course content and design approximates what a training course on a particular topic should be. This level of validity is sufficient if the course is to be used only for self-improvement. Content Validity requires that more than one SME review the content and certify that the content, the objectives, and the criterion test items are congruent. The readability of course content will be validated to ensure adherence to the identified level. Dick and Carey (1985) recommend using nine SMEs to establish content validity. Practically speaking, this places budgetary strain on the development of a project, so a minimum of three is recommended. Three, an odd number, will ensure that there is always a majority opinion and will negate the necessity of revisiting a content issue after the initial validity is conducted.

An in-house review of the course should be conducted to identify technical errors. Final review of the entire course will be conducted by the evaluation coordinator at the end of the Development Phase.

Small Group -- This evaluation is conducted at the end of the Development Phase with a small group of reviewers provided by the client. These reviewers should be additional Subject Matter Experts. The purpose of this activity is to determine the effectiveness of changes made following the one-to-one and in-house evaluations. If there will be no field test during Implementation, potential users from the representative audience should be included in this group to identify positive and negative indicators of course effectiveness. Reviewers in this evaluation are interviewed by the evaluation coordinator and are asked to complete survey forms and questionnaires.

Field Test -- This evaluation is conducted during Implementation in a situation as close to "real world" as possible. To conduct a field test use students from the target audience to test the effectiveness of the course on variables such as achievement, skill development, attitude, test item validity, and/or time spent on learning.

Summative Evaluation -- This evaluation is conducted when the training program is in use by the client. Summative evaluation, with its Experimental and Control group comparisons, will quantify the long-term effectiveness of the training program regarding knowledges, skills, behaviors, attitudes, or any other variables that the client established as the purpose for the course during Assessment/Analysis.

It is important to note that without at least content validity, a summative evaluation of a training program is in form only, producing neither reliable nor valid results. Test item validity is optimal and recommended as the best method of establishing the validity of a training program.

Phase IV -- Development involves the in-house review of the training program. Audio, video, graphics, and text are reviewed to identify and correct any major errors in the program. The result is a functional program that operates on the hardware system for which it was intended. The readability of the documentation is also conducted here to ensure adherence to identified reading levels established during Audience Analysis.

Phase V -- Implementation is the phase where a field test and summative evaluation are conducted. A pilot test is conducted that can be followed by a Summative Evaluation. A pilot test will identify minor errors in the program, identify areas of inconsistency and confusion for learners, and validate the tests used in the program. Positive results from a summative evaluation validate any changes made after the pilot study.

Summary

An evaluation plan expected to meet a project's needs must be well designed and congruent with the client's purpose for the course. The ultimate goal is to provide the client with a product that precisely matches the intended use. Evaluation must get more than just "lip-service" if a training company is to survive and remain competitive in the marketplace. Commitment of time and resources to product evaluation, in the form of quantitative and qualitative measures, ensures accountability and results in educationally sound products.

The rationale and components of evaluation must be explained to a client in non-technical terms emphasizing the return on money invested. Once clients are convinced of the importance, and see the results of thorough evaluation, the development company, the client, and the IVI industry as a whole will benefit. The development company benefits by demonstrating integrity in product development and thereby establishing a visible reputation for itself. The client benefits by having tangible proof that the resources spent achieved a return on investment. Finally, the industry benefits because the technology is solidly established as an effective instructional delivery system.

References

Basarab, D. (1989) . A practical approach to implementing a formal education evaluation process within the corporate world. Educational Technology, 29(11), p. 52-54.

Bloom, B. (Ed.). (1956). Taxonomy of educational objectives: Handbook 1. Cognitive domain. New York: David McKay.

Borg, W. & Gall, M. (1983). Educational research: An introduction. Fourth edition. New York: Longman.

Briggs, L. (Ed.). (1977). Instructional design: Principles and applications. Englewood Cliffs, NJ: Educational Technology Publications.

Dick, W. & Carey, L. (1985). The systematic design of instruction. Second edition. Glenview, Illinois: Scott, Foresman & Co.

Lee, William W. (1990). Evaluating interactive video training ensures instructional soundness. Performance & Instruction, 29(5), p.17-20.

Miller, A. & Irving, J. (1988). A pragmatic approach to evaluation of interactive media training programs. Thirtieth ADCIS Conference Proceedings. Philadelphia, PA., pp. 194-199.

Miller, R. (1989). The writing on the wall. Videodisc Monitor, 7(11), p. 2.

Reeves, T. (in press). The role, methods and worth of evaluation in instructional design. In K. Johnson and L. Foa (Eds.). Instructional design: New strategies for education and training. New York: MacMillan.